**Reanalysis of Simplex Correlations Support a Bifactor Representation**

Alexandre J. S. Morin, Concordia University
Joshua L. Howard, Monash University
Marylene Gagné, University of Western Australia

Cite as:
Morin, A.J.S., Howard, J.L., & Gagné, M. (2019). Webnote: Reanalysis of Simplex Correlations Support a Bifactor Representation. Montreal, QC: Substantive Methodological Synergy Research Laboratory.

Advocates of the simplex approach to represent the continuum structure of self-determined motivation (Sheldon et al., 2017) have argued that one factor approaches such as that implicit in Chemolli and Gagné (2014) Rash analyses of motivation were incompatible with a simplex representation of the structure of human motivation. More precisely, this claim is anchored on statistical evidence provided by McDonald (1980; also see Browne, 1982, 1992) showing that a simplex correlation matrix could best be reflected by two alternative two-factor models. Despite its intuitive appeal, this rationale is flawed for a variety of reasons. First, this rationale is based on a correlation matrix which, despite following a simplex structure, remains based on scale scores on subtests uncorrected for measurement errors. Second, the factor structures proposed by McDonald are highly artificial. Thus, one of those structures involves two orthogonal factors characterized by similar loadings associated with the variables located in the middle of the simplex, and loadings on the second factor that are a linear function of the loadings on the first factor. The second model proposed by McDonald, despite allowing the factors to be correlated, involves the estimation of a second factor in which the loadings associated with all variables are equal to one another. Third, our own re-analysis[1] showed that this correlation matrix (reported in Table 4 in McDonald, 1980, p. 174) could also be fitted to a one factor model with a generally satisfactory level of fit to the data (e.g., $\chi^2$ = 39.86, df = 9, p ≤ .01; CFI = .954; TLI = .923; SRMR = .038), although the RMSEA (.103) suggests possible misfit. This misfit could be partially due to the reliance on scale scores (we come back shortly to this statistical issue). More importantly, even though this solution was not able to achieve a level of fit comparable to that of McDonald's solutions, it resulted in interpretable factor loadings matching the

[1] All analyses reported in this webnote were realized with Mplus 8.0 Maximum Likelihood (ML) estimator (Muthén & Muthén, 2018).

rank ordering of the variables on the simplex ($\lambda$ = .618; .650; .729; .745; .714; .643).

Importantly, by conducting the re-analysis at the subtest level, it was not possible to clearly disaggregate the global versus subscale-specific variance present at the item level (as done by Howard et al., 2018 and Litalien et al., 2017). To address these limitations, we conducted a small simulation study in which a population model was generated including six correlated factors, derived from three indicators each, with no cross-loadings allowed between any of the factors, no correlated uniqueness, and no global factor. The factor loadings and uniquenesses used to simulate this data were directly taken from the first-order CFA solution reported by Howard et al. (2018) to match parameter estimates typically observed in SDT research. In contrast, the factor correlations were simulated to correspond to the same simplex-like correlation matrix (reported in Table 4 from McDonald, 1980, p. 174), using a sample size of 326 matching McDonald's specifications. From this population model, 1000 replication samples were generated and analyzed using a bifactor-CFA model. Interestingly, this model was found to fit the data quite well (across replications: $\chi^2$ = 137.322, df = 117, *ns*; RMSEA = .021; SRMR = .039)[2], and resulted in G-factor loadings matching the position of the indicators on the underlying continuum. Similar analyses conducted based on correlations matrices described by Guttman (1954) as corresponding to a hypothetical equally-spaced perfect simplex (see Table 2 in Guttman, 1954, p. 271: $\chi^2$ = 143.114, df = 75, $p \leq .01$; RMSEA = .052; SRMR = .066) or to a true correlation matrix obtained among six numerical ability tests (see Table 6 in Guttman, 1954, p. 284: $\chi^2$ = 167.295, df = 117, $p \leq .01$; RMSEA = .041; SRMR = .057). The average G-factor loadings obtained across replications from these three solutions are reported in Table 1. The simulation input and output files used in this demonstration are available at the end of this webnote.

Importantly, despite the fact that these factor loadings match the expected position of the indicators on the underlying simplex, differences are not as marked as those reported by Howard et al. (2018) and Litalien et al. (2017). Given the fact that these factor correlations match a known simplex matrix, this observation suggests that something more than a simplex structure characterizes motivation ratings and that this something requires bifactor modeling, or perhaps person-centered

[2] CFI and TLI estimates are not provided by Mplus 8.0 Monte Carlo simulation facilities.

analyses, to be correctly identified.

## References


Browne, M.W. (1982). Covariance Structures. In D.M. Hawkins (Ed), Topics in applied multivariate analysis (pp. 72-141). New York, NY: Cambridge University Press.

Browne, M.W. (1992). Circumplex models for correlation matrices. *Psychometrika, 57,* 469–497.

Chemolli, E., & Gagné, M. (2014). Evidence against the continuum structure underlying motivation measures derived from self-determination theory. *Psychological Assessment, 26*, 575–585.

Guttman, L. (1954). A new approach to factor analysis: the Radex. In Lazarsfeld, Paul. F. (ed.), *Mathematical thinking in the social sciences*. Glencoe, Ill: The Free Press.

Howard, J. L., Gagné, M., Morin, A. J. S., & Forest, J. (2018). Using bifactor exploratory structural equation modeling to test for a continuum structure of motivation. *Journal of Management*, *44*, 2638-2664.

Litalien, D., Morin, A. J. S., Gagné, M., Vallerand, R. J., Losier, G. F., Ryan. R. M. (2017). Evidence of a continuum structure of academic self-determination: A two-study test using a bifactor-ESEM representation of academic motivation. *Contemporary Educational Psychology*. *51*, 67-82.

Muthén, L.K., & Muthén, B. (2018). *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén.

Sheldon, K.M., Osin, E.N., Gordeeva, T.O., Suchkov, D.D., & Sychev, O.A. (2017). Evaluating the dimensionality of self-determination theory's relative autonomy continuum. *Personality & Social Psychology Bulletin*, *43*, 1215-1238.

**Table 1**

*Loadings Obtained on the Global Factor across Replications*

| Items | McDonald (1980): Table 4 | Guttman (1954): Table 2 | Guttman (1954): Table 6 |
|---|---|---|---|
| Factor 1 Item 1 | .525 | .482 | .568 |
| Factor 1 Item 2 | .554 | .506 | .599 |
| Factor 1 Item 3 | .563 | .518 | .610 |
| Factor 2 Item 1 | .417 | .499 | .465 |
| Factor 2 Item 2 | .536 | .643 | .599 |
| Factor 2 Item 3 | .525 | .628 | .585 |
| Factor 3 Item 1 | .422 | .442 | .487 |
| Factor 3 Item 2 | .504 | .527 | .582 |
| Factor 3 Item 3 | .434 | .452 | .498 |
| Factor 4 Item 1 | .390 | .312 | .417 |
| Factor 4 Item 2 | .534 | .420 | .576 |
| Factor 4 Item 3 | .457 | .361 | .494 |
| Factor 5 Item 1 | .467 | .271 | .413 |
| Factor 5 Item 2 | .456 | .264 | .404 |
| Factor 5 Item 3 | .470 | .272 | .416 |
| Factor 6 Item 1 | .460 | .482 | .447 |
| Factor 6 Item 2 | .410 | .506 | .399 |
| Factor 6 Item 3 | .446 | .518 | .436 |

**Syntax based on McDonald (1980), Table 4.**

```
Title: Data generation input;
montecarlo:
        names =  a1-a3 b1-b3 c1-c3 d1-d3 e1-e3 g1-g3;
        ngroups = 1;
        nobs = 326;
        nreps = 1000;
model population:
! Main loadings
F1 BY a1*.84;
F1 BY a2*.88;
F1 BY a3*.90;
F2 BY b1*.63;
F2 BY b2*.81;
F2 BY b3*.79;
F3 BY c1*.58;
F3 BY c2*.69;
F3 BY c3*.59;
F4 BY d1*.53;
F4 BY d2*.72;
F4 BY d3*.62;
F5 BY e1*.67;
F5 BY e2*.65;
F5 BY e3*.67;
F6 BY g1*.73;
F6 BY g2*.65;
F6 BY g3*.71;

! Intercepts
[a1-g3@0];

!Uniquenesses
a1*.29;
a2*.23;
a3*.19;
b1*.60;
b2*.34;
b3*.38;
c1*.67;
c2*.53;
c3*.66;
d1*.72;
d2*.49;
d3*.62;
e1*.55;
e2*.58;
e3*.55;
g1*.47;
g2*.58;
g3*.50;

!Latent means
[f1@0];
[f2@0];
[f3@0];
[f4@0];
[f5@0];
[f6@0];
```

```
!Latent variances and covariances
f1@1;
f2@1;
f3@1;
f4@1;
f5@1;
f6@1;
f1 WITH f2*.51;
f1 WITH f3*.46;
f1 WITH f4*.46;
f1 WITH f5*.40;
f1 WITH f6*.33;
f2 WITH f3*.51;
f2 WITH f4*.47;
f2 WITH f5*.39;
f2 WITH f6*.39;
f3 WITH f4*.54;
f3 WITH f5*.49;
f3 WITH f6*.47;
f4 WITH f5*.57;
f4 WITH f6*.45;
f5 WITH f6*.56;

MODEL :
F1 BY a1* a2-a3;
F2 BY b1* b2-b3;
F3 BY c1* c2-c3;
F4 BY d1* d2-d3;
F5 BY e1* e2-e3;
F6 BY g1* g2-g3;
GF BY a1* a2-g3;
F1-F6@1;
GF@1;
GF WITH F1-F6@0;
F1 WITH F2-F6@0;
F2 WITH F3-F6@0;
F3 WITH F4-F6@0;
F4 WITH F5-F6@0;
F5 WITH F6@0;
```

**Syntax based on Guttman (1954), Table 2.**

```
Title: Data generation input;
montecarlo:
        names =  a1-a3 b1-b3 c1-c3 d1-d3 e1-e3 ;
        ngroups = 1;
        nobs = 326;
        nreps = 1000;
model population:
! Main loadings
F1 BY a1*.84;
F1 BY a2*.88;
F1 BY a3*.90;
F2 BY b1*.63;
F2 BY b2*.81;
F2 BY b3*.79;
F3 BY c1*.58;
F3 BY c2*.69;
F3 BY c3*.59;
F4 BY d1*.53;
F4 BY d2*.72;
F4 BY d3*.62;
F5 BY e1*.67;
F5 BY e2*.65;
F5 BY e3*.67;

! Intercepts
[a1-e3@0];

!Uniquenesses
a1*.29;
a2*.23;
a3*.19;
b1*.60;
b2*.34;
b3*.38;
c1*.67;
c2*.53;
c3*.66;
d1*.72;
d2*.49;
d3*.62;
e1*.55;
e2*.58;
e3*.55;

!Latent means
[f1@0];
[f2@0];
[f3@0];
[f4@0];
[f5@0];

!Latent variances and covariances
f1@1;
f2@1;
f3@1;
f4@1;
f5@1;
```

```
f1 WITH f2*.60;
f1 WITH f3*.36;
f1 WITH f4*.216;
f1 WITH f5*.1296;
f2 WITH f3*.6;
f2 WITH f4*.36;
f2 WITH f5*.216;
f3 WITH f4*.6;
f3 WITH f5*.36;
f4 WITH f5*.60;

MODEL :
F1 BY a1* a2-a3;
F2 BY b1* b2-b3;
F3 BY c1* c2-c3;
F4 BY d1* d2-d3;
F5 BY e1* e2-e3;
GF BY a1* a2-e3;
F1-F5@1;
GF@1;
GF WITH F1-F5@0;
F1 WITH F2-F5@0;
F2 WITH F3-F5@0;
F3 WITH F4-F5@0;
F4 WITH F5@0;
```

**Syntax based on Guttman (1954), Table 6.**

```
Title: Data generation input;
montecarlo:
        names =  a1-a3 b1-b3 c1-c3 d1-d3 e1-e3 g1-g3;
        ngroups = 1;
        nobs = 240;
        nreps = 1000;
model population:
! Main loadings
F1 BY a1*.84;
F1 BY a2*.88;
F1 BY a3*.90;
F2 BY b1*.63;
F2 BY b2*.81;
F2 BY b3*.79;
F3 BY c1*.58;
F3 BY c2*.69;
F3 BY c3*.59;
F4 BY d1*.53;
F4 BY d2*.72;
F4 BY d3*.62;
F5 BY e1*.67;
F5 BY e2*.65;
F5 BY e3*.67;
F6 BY g1*.73;
F6 BY g2*.65;
F6 BY g3*.71;

! Intercepts
[a1-g3@0];

!Uniquenesses
a1*.29;
a2*.23;
a3*.19;
b1*.60;
b2*.34;
b3*.38;
c1*.67;
c2*.53;
c3*.66;
d1*.72;
d2*.49;
d3*.62;
e1*.55;
e2*.58;
e3*.55;
g1*.47;
g2*.58;
g3*.50;

!Latent means
[f1@0];
[f2@0];
[f3@0];
[f4@0];
[f5@0];
[f6@0];
```

```
!Latent variances and covariances
f1@1;
f2@1;
f3@1;
f4@1;
f5@1;
f6@1;
f1 WITH f2*.62;
f1 WITH f3*.62;
f1 WITH f4*.54;
f1 WITH f5*.29;
f1 WITH f6*.28;
f2 WITH f3*.67;
f2 WITH f4*.53;
f2 WITH f5*.38;
f2 WITH f6*.37;
f3 WITH f4*.62;
f3 WITH f5*.48;
f3 WITH f6*.52;
f4 WITH f5*.62;
f4 WITH f6*.57;
f5 WITH f6*.64;

MODEL :
F1 BY a1* a2-a3;
F2 BY b1* b2-b3;
F3 BY c1* c2-c3;
F4 BY d1* d2-d3;
F5 BY e1* e2-e3;
F6 BY g1* g2-g3;
GF BY a1* a2-g3;
F1-F6@1;
GF@1;
GF WITH F1-F6@0;
F1 WITH F2-F6@0;
F2 WITH F3-F6@0;
F3 WITH F4-F6@0;
F4 WITH F5-F6@0;
F5 WITH F6@0;
```